CERES Data Management Activity

Presented to CERES Science Team Meeting

April 26, 2011

Jonathan Gleason

Jonathan.L.Gleason@nasa.gov

Data Management Team Members

Instrument:
Denise Cooper
Thomas Grepiotis
Richard Spivak
Mark Timcoe
Nelson Hillyer
Dianne Snyder

Production / Optimization:
Lisa Coleman
Carla Grune
Brian Magill

ERBElike:
Dale Walikainen
Jeremy Lande

Convolution:
Walter Miller
Igor Antropov

Inversion:
Victor Sothcott

SSAI:
Lee Bodden

TISA Gridding:
Raja Raju
Hazari Syed

TISA Averaging:
Cathy Nguyen
Betty Lock

SARB:
Tom Caldwell

FM5:
Jim Closs

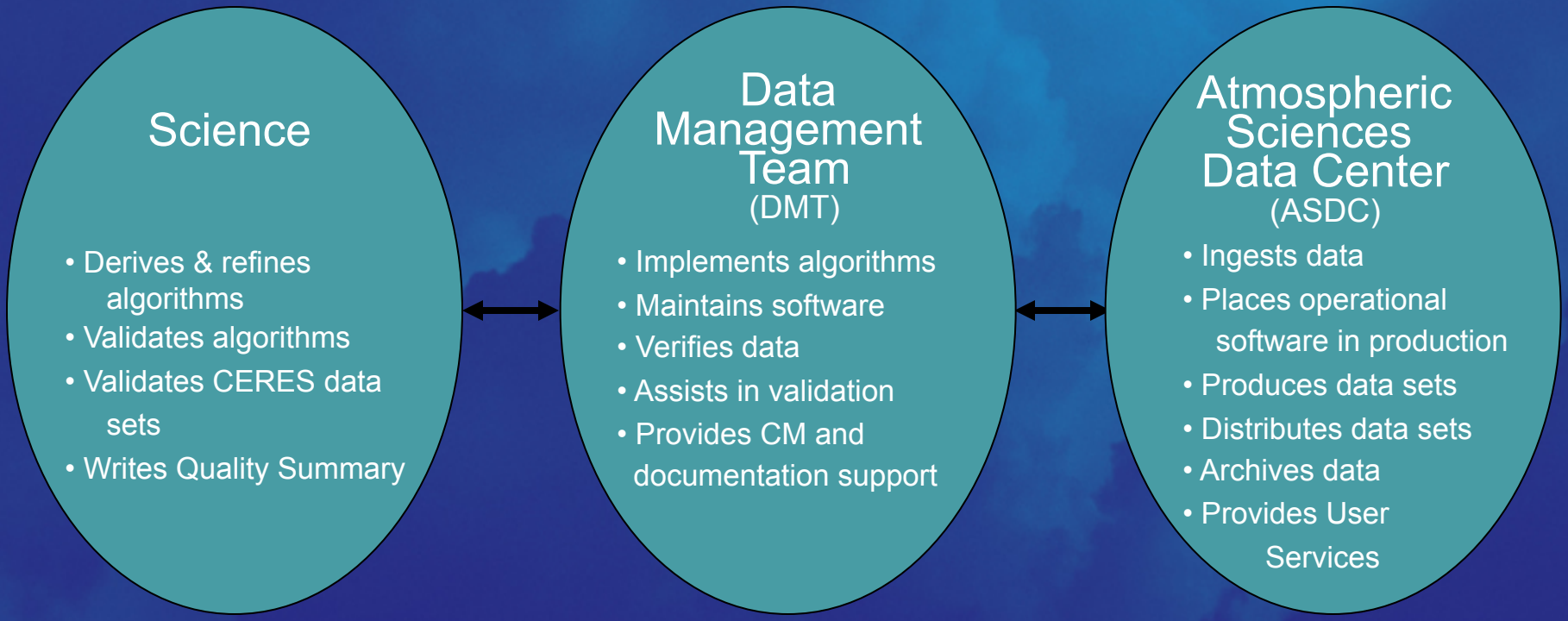
Clouds:
Sunny Sun-Mack
Ricky Brown
Yan Chen
Liz Heckert
Rita Smith
Sharon Gibson

Configuration Management:
Tammy Ayers
Joanne Saunders

Topics To Be Covered

- DMT overview
- DMT Activity since last STM
- Edition 2 and 3 data availability
- Production schedule
- NPP status
- GMAO GEOS-5 update
- DMT/ASDC Process Improvement

CERES Organization



CERES Subsystems

- CERES is made up of 7 Working Groups
  - Instrument
  - ERBElike
  - Clouds
  - Inversion or ADM
  - SOFA
  - SARB
  - TISA
- Code organized into 12 Subsystems
  - Each subsystem tied to 1 or more working groups
- Each Subsystem made up of 1 or more Product Generation Executives (PGEs)
  - Currently there are 78 active PGEs

CERES Processing Software

Subsystem Number	Subsystem Name	LOC (to nearest 1K)	Publicly Available Data Products	Product Frequency	Comments
	CERESlib	133K*			All Satellites
1	Instrument/Pre-Processor	4K			NPP only
1	Instrument	306K	BDS	1/day	All Satellites
2	ERBE-like/ Inversion	30K	ES-8	1/day	All Satellites
3	ERBE-like/ TSA	12K	ES-9, ES-4	1/month	All Satellites
12	MOA	30K*			Run monthly
4.1 – 4.4	Clouds	503K*			All Satellites
4.5 – 4.6	Inversion	227K	SSF	1/hour	All Satellites
5	SARB	164K	CRS	1/hour	All Satellites
6 & 9	TISA-Gridding	60K	FSW, SFC, ISCCP-D2like-Day/Nit	60/month, 36/month, 1/month	All Satellites
11	GGEO	172K	ISCCP-D2like-GEO	1/month	Geostationary
7.2	Synoptic SARB	47K			All Satellites
7.1 & 8 10	TISA-Averaging	249K	SYN, AVG, ZAVG SRBAVG	1/day, 1/month, 1/month 5/month	All Satellites
	TOTAL LOC	1,959K			

Activity since Fall STM

31 total deliveries since 9/16/2010

- Instrument
  - 1 Ada to AMI-P, 3 C++, 2 Delta
- Clouds
  - 1 Beta Edition 4
- Inversion
  - 2 Ed3 fixes, 3 Ed1-CV
- Tisa Grid
  - 1 ISCCP d2like fix, 2 Ed3 LUTs
- Tisa Averaging
  - 3 new Ed3, 1 Ed 2 fix
- Inst SARB
  - 1 MATCH delivery, 1 Ed2 fix
- Synoptic SARB
  - 1 AMI-P conversion, 1 Ed2/3 fix
- CERESlib
  - 1 Ed4 support, 1 submission scripts
- PerlLib
  - 4 submission scripts
- GGEO
  - 2 delivery offline produced files

Edition 2 Data Availability

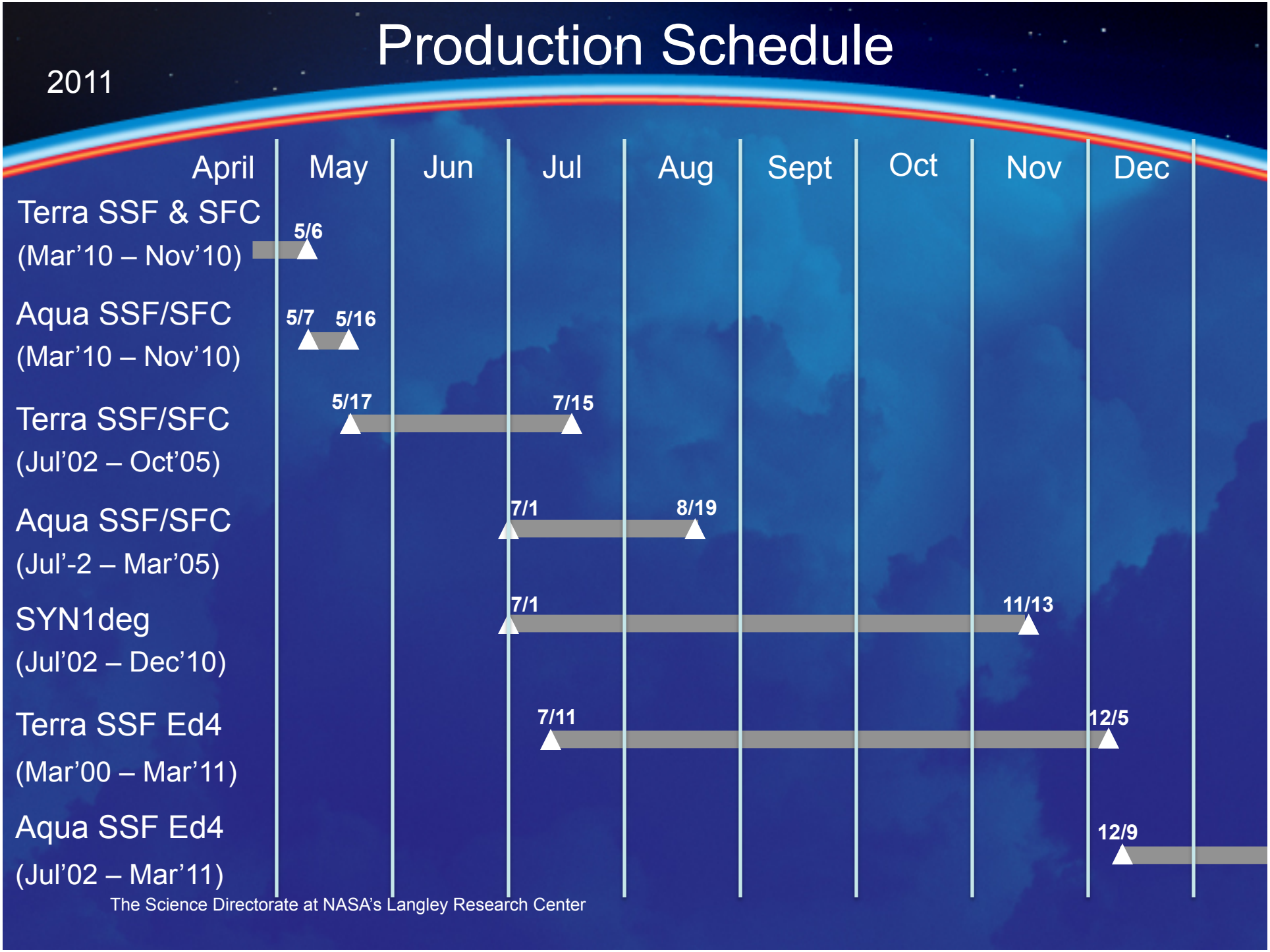
Product	Edition 2	Available through	Comments
BDS, ES8, ES9, ES4	Edition2 (Terra, Aqua)	June ‘10	No more expected
SSF	Edition2G (Terra) Edition2D (Aqua)	June ‘10	
SFC	Edition2G (Terra) Edition2D (Aqua)	June ‘10	
SRBAVG	Edition2D (Terra) Edition2A (Aqua)	Oct ’05	
CRS	Edition2G (Terra) Edition2C (Aqua)	June ’10 Dec ’07	
FSW	Edition2G (Terra) Edition2C (Aqua)	Feb ’10 Dec ‘07	
SYN, AVG, ZAVG	Edition2C (Terra) Edition2B (Aqua)	Oct ’05	

Edition 3 Data Availability

Product	Edition 3	Available through	Comments
BDS	Terra & Aqua	Dec 1, 2010	
SSF	Terra Aqua	10/31/2005 – 6/1/2010 4/1/2005 – 3/1/2010	Currently Processing
SFC	Terra Aqua	Nov '05 – Feb '10 April '05 – Feb '10	Currently Processing
SYN1deg	Terra + Aqua merged	Pending	Processing start 7/1/2011
SSF Edition 1-CV	Terra Aqua	7/1/2010 – 1/1/2011	Internal Product

Production Schedule

2011

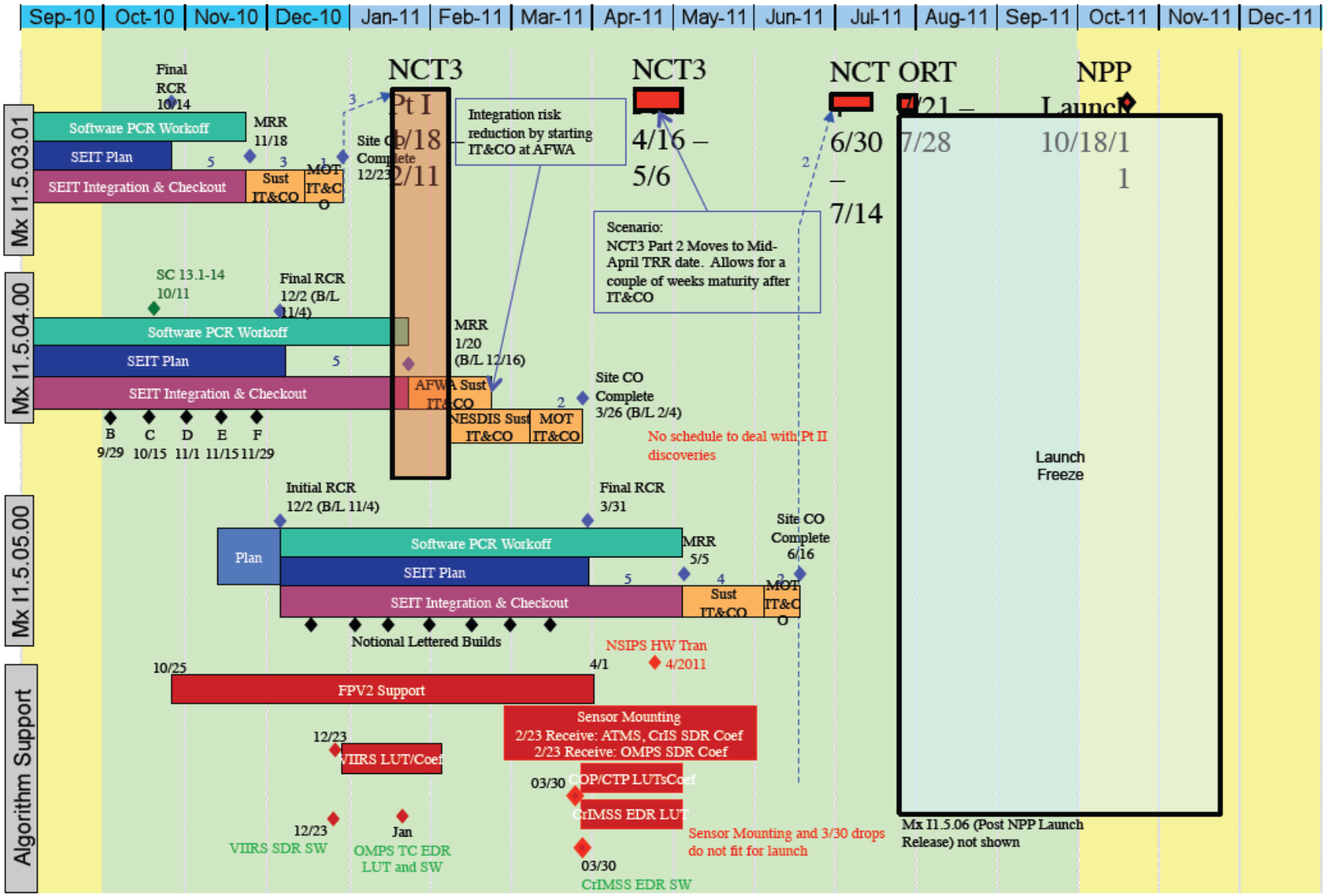


The Science Directorate at NASA's Langley Research Center

CERES on NPP

- For NPP CERES leverages existing network infrastructure and will ingest all data from GSFC Land PEATE
  - CERES science, diagnostic, and telemetry RDRs
  - VIIRS subsampled radiance data
  - VIIRS AOT
- Instrument preprocessor developed to convert NPP data to EOS format (no other code change *required*)
- Converted Ada code to C++
- Developed, delivered, and verified code to subset VIIRS RDRs at Land PEATE

IDPS Sustainment Schedule v24-10/12/2010



NPP Confidence Test 3

- Two Part Test, Ground system mostly part 2 (RFR April 12th – 14th)
- Primary goal to test end-to-end data flow from downlink station to Science teams
  - Flowed 44 continuous orbits (72 hours) of proxy and real spacecraft sensor data
  - ASDC received:
    - CERES RDRs (received 375 of about 400)
    - VIIRS subsetted radiance data
    - VIIRS subsetter input (VIME, VMAE, VDNE)
- Ran Instrument NPP Preprocessor and main processor
  - Geolocation correction in Preprocessor
  - Filename format correction in Preprocessor
- Preliminary comparison of subsetted VIIRs files good

FM5 Code Development

- Instrument only subsystem modified
  - Preprocessor delivered to AMI-P and supported NCT3 test (implementing fix)
  - Ada code already promoted on AMI-P (pending memory dump byte swap fix – deliver 5/6)
  - C++ main processor promoted (implementing fix)
- 13 total PGEs delivered
  - 8 Ada (Production expected July 1, 2011)
  - 5 C++ PGEs (ready to replace Ada Oct 2011)
- VIIRs subsetter updates delivered to Land PEATE and test output from Land PEATE verified

GMAO G5 Data Sets

- G5.2
  - Used for Edition 2 and Edition 3 processing
  - Availability: Dec 2007 – current
  - End stream no earlier than Mar 2012
- G5.4
  - Used for Edition 4 processing
  - Now reprocessing Dec 1997 – current
  - Currently have 26 months, stream = 11x
  - Model and Analysis improvements
- G5.7.1
  - Used for FLASHflux only
  - Forward Processing only, production to begin May 24
  - Horizontal resolution increase, temporal resolution increase, new variables
  - HDF-4 to netCDF-4/HDF-5

Process Optimization Effort

- Migration from legacy SGI to AMI Cluster
  - Job submission script development
- Production Request Database
- PGE exit code standardization
- Pre-delivery robust testing procedures
- SCCR Dashboard database

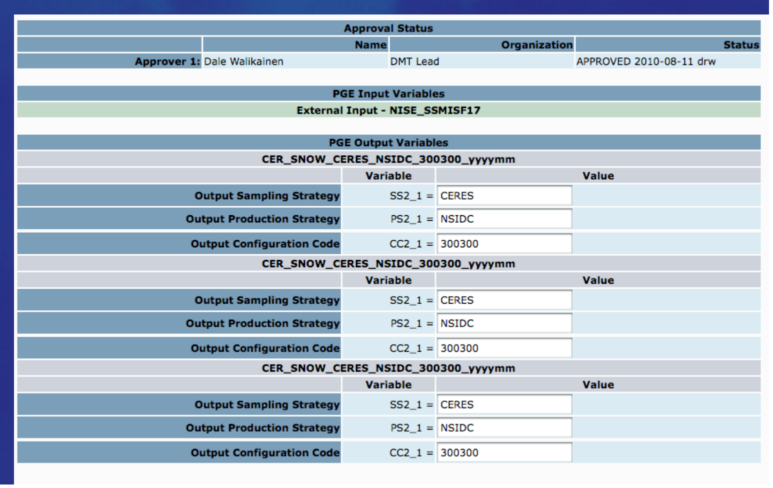
Computing Platform Migration

- FLASHflux end-to-end AMI-P test late May
  - 1 week of concurrent AMI-P/Warlock production
- Instrument production on AMI-P start July 1, 2011
  - 1 month of concurrent AMI-P/Warlock production
- Turn off Warlock code mid-August

Warlock (SGI)	**Magneto (P4)**	**AMI-P (P6 & x86)**
• Instrument • FLASHflux	• ERBElike • Clouds (Ed2) • Inversion (Ed2 & Ed3) • TISA Gridding (Ed2 & Ed3) • Inst. SARB (Ed2) • Synoptic SARB (Ed2 & Ed3) • TISA Averaging (Ed2) • GGEO • MOA (Ed2)	• Instrument • Clouds (Ed4 Beta) •TISA Averaging (Ed3) • MOA (Ed4) In Development • ERBElike • Inversion (Ed4 Beta) • Synoptic SARB (Ed3) • ISCCP-D2like • FLASHflux

Production Request Database

- Production Request Database Tool
  - PRs submitted, reviewed, updated and approved via web interface
  - Public view and Restricted write
  - Capability for scripting to automate production (requires standardized PGE exit codes)
- Prototype demonstrated January 2010
- Build 1 August 2011

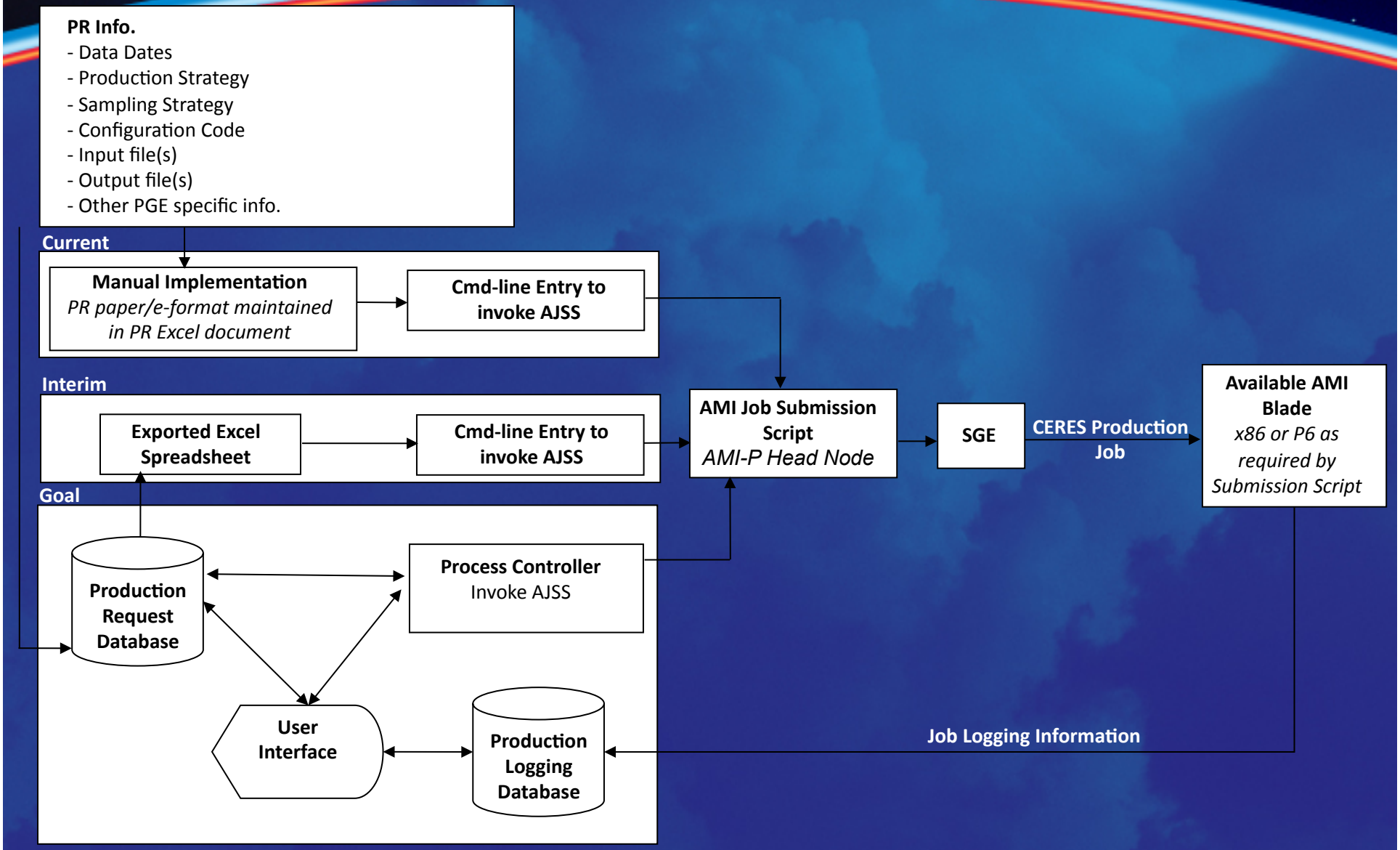


Approval Status			
	Name	Organization	Status
Approver 1:	Dale Walikainen	DMT Lead	APPROVED 2010-08-11 drw

PGE Input Variables
External Input - NISE_SSMISF17

PGE Output Variables		
CER_SNOW_CERES_NSIDC_300300_yyyymm		
	Variable	Value
Output Sampling Strategy	SS2_1 =	CERES
Output Production Strategy	PS2_1 =	NSIDC
Output Configuration Code	CC2_1 =	300300
CER_SNOW_CERES_NSIDC_300300_yyyymm		
	Variable	Value
Output Sampling Strategy	SS2_1 =	CERES
Output Production Strategy	PS2_1 =	NSIDC
Output Configuration Code	CC2_1 =	300300
CER_SNOW_CERES_NSIDC_300300_yyyymm		
	Variable	Value
Output Sampling Strategy	SS2_1 =	CERES
Output Production Strategy	PS2_1 =	NSIDC
Output Configuration Code	CC2_1 =	300300

PR Retrieval Automation On AMI



Robust Pre-Delivery Testing

- Kaizen team recommended more testing “upstream” of software delivery
- Identify specific test cases for each PGE
  - Run with all data
  - Run with less than minimum required and bad data
  - Run full month or year
- ERBElike delivery to AMI-P first subsystem
  - Null case and full month testing have already identified problems

SCCR Dashboard Database

- Online status updates
- Track dates for:
  - Requirement issuance
  - Development
  - Testing
  - Delivery
  - Promotion
  - Data public release
- Updated biweekly for lessons learned

SCCR Dashboard

SCCR Status-Update | Search for SCCR | Assign Priority | CPOB SCCR Update | SCCR Status Update | Create SCCR | Edit SCCR | View Subsystems | View Platforms

SCCR Status-Update

Search:

Pro. Prio.	SCCR	Subsystem	Dep.	Reason for Delivery	Platforms	Last Completed Process Step	Completion Date	Current Process Step	Status from Last Meeting	Current Status
1	842	Instrument	-	Fix issues discovered in NCT3 part 2.	AMI-x86 AMI-P6	Create SCCR	03/23/2011	Finalize SCCR	May 6 delivery date.	[Denise Cooper] - A fix for the issues discovered during NCT 3 Part II pre-test has been made and is being tested before a delivery after NCT 3 Part II is complete. (04/12/2011)
2	843	Instrument	-	Fix problem with Gimbal and Memory Dump data output from CER1.1P1,3,5, & 7.	AMI-x86	Create SCCR	03/28/2011	Finalize SCCR	No updates.	[Denise Cooper] - A possible fix for the memory dump byte ordering has been made and is being tested. (04/12/2011)
3	838	Instrument	-	Correction for the radiance and double drift corrected count values.	AMI-P6	Create SCCR	03/16/2011	Finalize SCCR	No updates.	-
				Initial delivery of all	AMI-P6	Release to		Perform	Summary file to be added to this; shooting for	[Tammy Ayers] - Updated code has been redelivered to CM. On hold

Showing 1 to 20 of 20 entries

Generate PDF SCCR Status-Update Report
Generate PDF SCCR Priority Report
Generate PDF SCCR Creation Report
Generate PDF SCCR Promotion Report

SCCR Dashboard Prototype
Version 0.1 (alpha)
thomas.n.hillyer@nasa.gov

SCCR 716

Technical Information

Subsystem	Open	PGE(s)	New PGE(s)	Needs CERESlib Update	Certified Platforms	Description	Last Modification Timestamp
Instrument	X	1.1P1, 1.1P3, 1.1P5, 1.2P5, 1.3P1, 1.3P2, 1.3P3	X	X	AMI-P6 AMI-x86	Ada Terra & Aqua AMI-P	2011-02-01 21:20:20

Index	Date	Originator	Group	Comment
6	04/12/2011	Denise Cooper	Subsystem	ValR17s are being evaluated. A problem with the comparison of early Aqua Edition1-CV vs. ValR17 has been uncovered. Attempting to determine which set of data has the issue, Edition1-CV or ValR17. It appears that not all the data was staged for one of the runs.
5	03/31/2011	Denise Cooper	Subsystem	This SCCR is closed. ValR17s are being evaluated.
4	03/31/2011	Tonya Davenport	ASDC	Promoted 3/18/11
3	03/31/2011	Angel Cross	ASDC	Continuing to test against epilogue.
2	03/15/2011	Tammy Ayers	CM	Provided updated Instrument (SCCR 716) script to the ASDC. The update allows the platform to be specified.
1	03/15/2011	Denise Cooper	Subsystem	SGE Driver for CER1.2P1 was updated to include the platform parameter. The script was sent to SIT so they could complete their testing.

Process Tracking

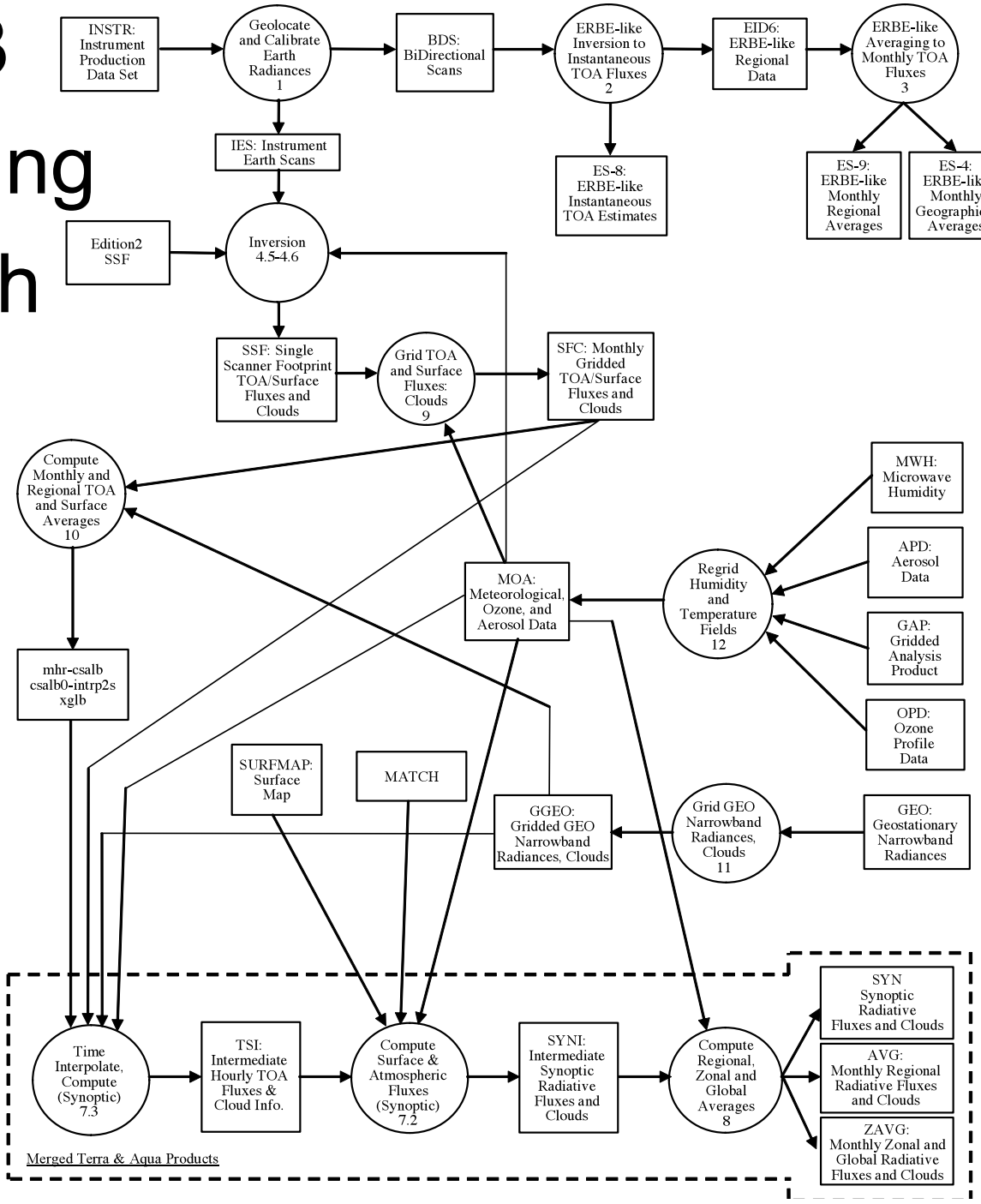
Responsible Team	Step	Date
Subsystem	Create SCCR	2009-07-02
	Finalize SCCR	
	Preliminary Delivery Memo to CM	
	Update Code & Unit Test	
	Integrate Code & Robust Testing	
	Prepare Delivery Package	
	Delivery to CERES CM	2011-01-27
Configuration Management	Test According to Test Plan	
	Release to Langley DAAC	2011-01-27
ASDC	Perform Operational Testing in PPE	

Questions & Comments

View and Order CERES Data

http://ceres.larc.nasa.gov/order_data.php

Edition 3 Processing Approach



Data from other Instruments used by CERES

- CERES Instrument/ERBElike only subsystems that can process when only CERES data available
- CERES directly uses the following MODIS data sets:
  - MYD02SS1/MOD02SS1* (19 channel radiance subset of every other pixel every other scanline)
  - MYD03/MOD03* (geolocation)
  - MYD04_L2/MOD04_L2 (5 min 10 km aerosol swath)
  - MYD08_D3/ MOD08_D3 (daily 1 deg aerosol)
    - Critical data sets; must have matched pairs to process
- Additionally CERES uses Geostationary satellite data:
  - MET-5, MET-6, MET-7, MET-8, MET-9
  - GOES-8, GOES-9, GOES-10, GOES-11, GOES-12
  - GMS-5, MTSAT-1R

Recommendations

- Need requirements freeze “handshake”
  - Minimize rework in testing and delivery
- Verification testing to DMT
  - Identify robust testing per PGE
  - Much currently done at ASDC
- Pre Production Environment (PPE) testing at ASDC
  - In Place of ValRx
  - SCCR remains open (code fixes easier)
- PR, Control Func. and Error tracking Kaizens

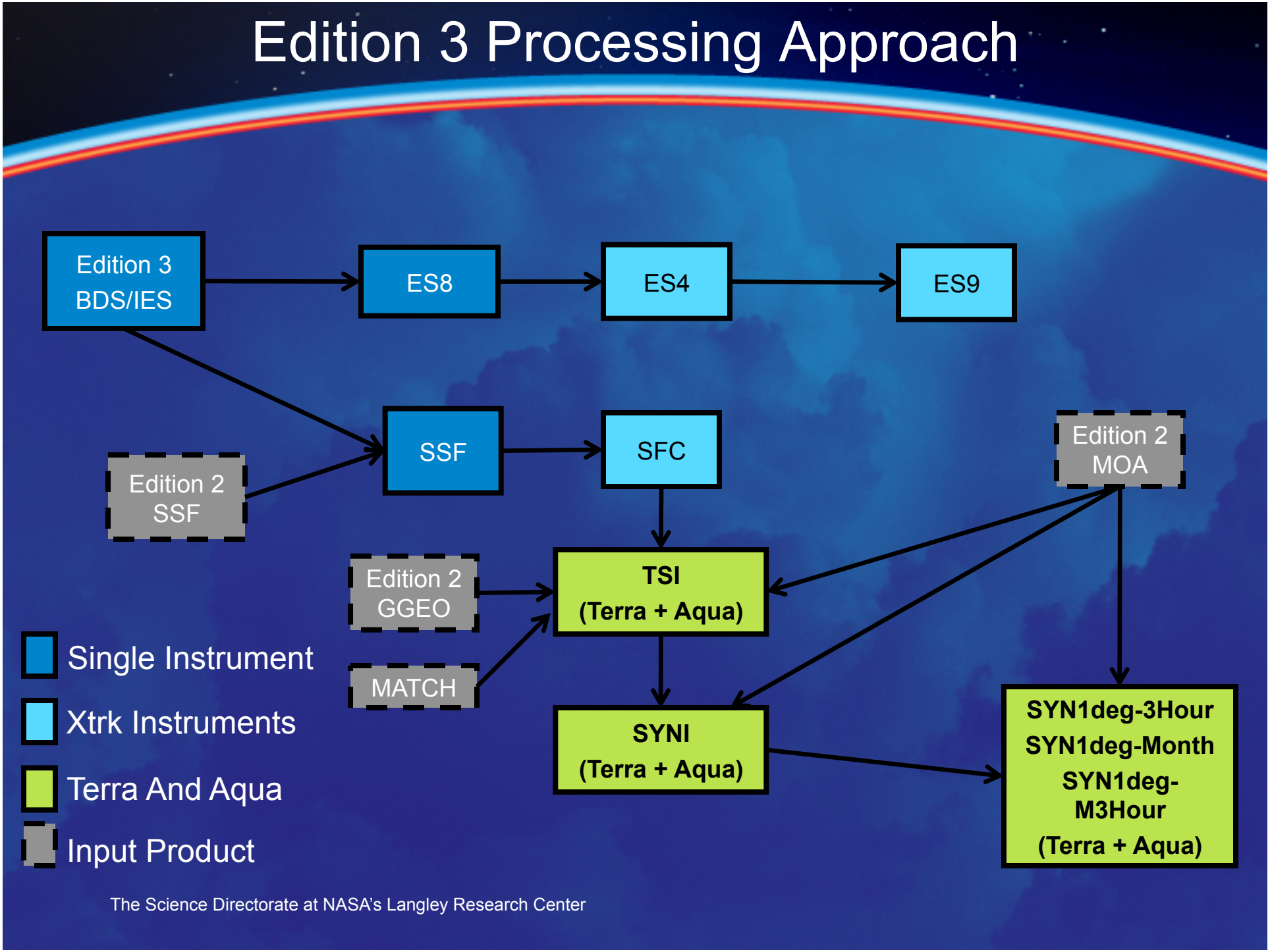
Production Request Database

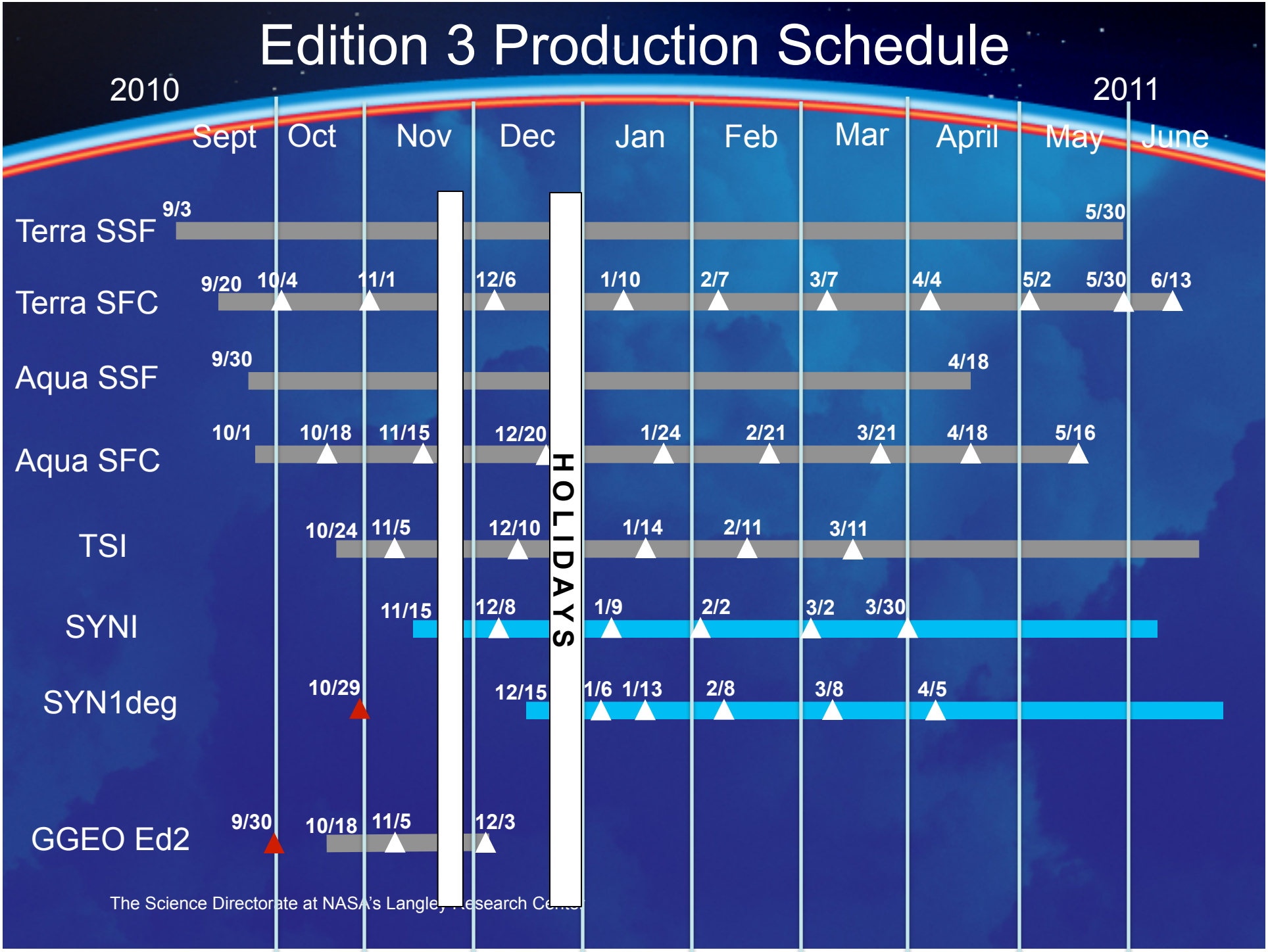
- PR requests a specified range of data dates be processed for a CERES PGE in the production environment
- Provides data dates, input/output file names, configuration code, target platform, PGE, etc
- Used for operational testing and production

Example PR

PR Year & Item_#	PGEs	Instrument (INST)	Input Production Strategy	Output Production Strategy	Begin Datadate to process	End Datadate to process	DAAC Verification (If no CC # provided, use most recent)
6/22/10 – Process Surface Albedo History maps for Edition3 Synoptic SARB Processing (SCCR 794) See note prior to PR 88-10 regarding the Edition3 CC number strategy **Computing Platform: magneto P4** **Environment/runtime variables:** **SS5_MATCH = C4 (dates up through June 2006)** **SS5_MATCH = C5 (dates beginning with July 2006 and forward)**							
104-10	5.0P2	FM1, FM2	PS12=DAO-GEOS4 PS4_5=Edition3A	PS5=Edition3A	7/2/02	2/28/06	CC4_5=300300 CC12= 016023 = 016024 (2/04) = 016025 (4/04) = 016026 (12/04) = 017027 (1/05) = 017028 (12/05) = 018029 (2/06) CC5=300300
103-10	5.0P2	FM1	PS12= DAO-GEOS4 PS4_5=Edition3A	PS5= Edition3A	3/1/2006	6/30/2006	cc4_5=300300 cc12= 018029 cc5=300300
102-10	5.0P2	FM1	PS12=DAO-GEOS4 PS4_5=Edition3A	PS5=Edition3A	7/1/2006	12/31/2007	cc4_5=301300 cc12=018029 =018030 (2/07) cc5=301300

Edition 3 Processing Approach




Edition 3 Production Schedule
2010
2011
Sept
Oct
Nov
Dec
Jan
Feb
Mar
April
May
June
HOLIDAYS
Terra SSF
9/3
5/30
Terra SFC
9/20
10/4
11/1
12/6
1/10
2/7
3/7
4/4
5/2
5/30
6/13
Aqua SSF
9/30
4/18
Aqua SFC
10/1
10/18
11/15
12/20
1/24
2/21
3/21
4/18
5/16
TSI
10/24
11/5
12/10
1/14
2/11
3/11
SYNI
11/15
12/8
1/9
2/2
3/2
3/30
SYN1deg
10/29
12/15
1/6
1/13
2/8
3/8
4/5
GGEO Ed2
9/30
10/18
11/5
12/3
The Science Directorate at NASA's Langley Research Center